



Natural Language Processing Research

FOR HUNTSVILLE.AI

MARCH 8, 2023

SCOTT B. TASHAKKOR, P.E. (NASA/MSFC)

Preface

- ▶ NASA operates in the Government Cloud and FedRamp approved environment
- ▶ Features are not always added
- ▶ Features are delayed
- ▶ Need to have isolated (firewalled) systems

Documentation creation/identification

- ▶ Advise and address common needs in documents
- ▶ Identification of language (contracts)
- ▶ Identify and highlight key data
 - ▶ Dates, prices
- ▶ Open formats
 - ▶ Invoices
 - ▶ Letters
 - ▶ Standards

Document consistency

- ▶ Documents can be inconsistent within themselves
- ▶ Numbers, dates, specifications
- ▶ This system grows exponentially
 - ▶ Each sentence not only inside the document but then document to document
 - ▶ Individual project definitions
- ▶ Open formats
 - ▶ Across agencies

Translation of technical documents

- ▶ Scientific papers written in other languages
 - ▶ Scientists will recognize equations or maybe abstract translated
- ▶ Identification
 - ▶ Figures will sometimes have symbols or systems will interpret
 - ▶ Language in the figures
 - ▶ Math equation identification
 - ▶ Greek letters and matching between references in paper

Speech to Text with analysis

- ▶ Want real-time speech translation to text
- ▶ Includes (near) real-time sentiment analysis
- ▶ Needs to be done with low size weight and power (SWaP)
- ▶ Different accents
- ▶ Technical terminology/jargon

Summary

- ▶ Just a sampling of some of the NLP projects
- ▶ Would prefer to utilize commercial products
- ▶ Have different challenges